UCD GEARY INSTITUTE DISCUSSION PAPER SERIES

A Model of Intervention in Childhood

Dr. James Symons
Claremont McKenna College

jsv_symons@yahoo.co.uk

Dr. Leon Feinstein
Institute for Education, Centre for Research on the Wider Benefits of Learning,

20 Bedford Way, London WC1H OAL, England

Associate at the Centre for the Economics of Education, London School of

Economics, England

lf@bg.ioe.ac.uk

Version: 2.3

27th March 2008

Abstract

This paper describes a model and resulting simulations to assess the appropriate age structure of intervention in childhood on the theme: should we intervene early or late? We use asset theory approaches to construct a general model of state investment whose aim is to reduce inequality in human capital. We set out the key parameters of such a model, clarifying the assumptions that must be made by state planners or economists in assessing the relative value of targeted investment at different ages in the presence of a range of elements of uncertainty. We simulate the model showing how the age-investment schedule will vary under different assumptions. Early investment is highly valued because of the likely decline with age in effectiveness but the trade-offs are strongly moderated by other important assumptions around which there is uncertainty or are choice variables of the state planner.

1 Introduction

There has been considerable debate amongst economists and policy experts as to the optimal timing of intervention in childhood to reduce adult social exclusion and enhance equity (see, for example, Heckman, 2003; Piatt, 2003; Reynolds, Wang & Walberg, 2003.) The consensus of much recent debate is that early is better and that delay in the support for those most in need makes it harder to reduce attainment gaps. In most countries, state expenditure on the early years is much below that on secondary age or college students. It is plausible that this is an inefficient use of public resources, whether the goal is to maximize average achievement in the population (efficiency) or reduce the disadvantage of those whose family background and personal circumstances leave them trailing better-off children even before they start school (equity).

This paper does not provide new evidence on the effectiveness or otherwise of early intervention. We accept that early intervention may be more effective than later intervention because of key neurological moments in early development (Goswami, 2007; Whitebread, 2007) and because of impacts on young people's motivation and sense of their own capabilities (Bandura, 2004; Wigfield & Eccles, 2000) i.e. success breeds success (Heckman *op. cit.*).

However there are important trade-offs in a system of targeted intervention between the benefits of providing early, preventive support and the cost associated with false identification of need. In earlier work (Feinstein, 2006), one of us analysed two UK Birth Cohorts (the National Child Development Study and British Cohort Study) showing that it is important to maintain a balanced perspective on risk continuity and discontinuity. Although for many children, educational under-performance carries long-term and profound sequelae, there is also considerable instability of risk. Many of those at risk at one age are not at risk at subsequent ages.

It follows that any efficient system of intervention must recognise not only that an early, preventive intervention will on the whole be more efficient than a later and larger reactive intervention, but also that targeting should be able to recognise and respond to changes in development and changes in risk. Against the likely greater effectiveness of early intervention must be placed the likely greater accuracy of later targeted intervention. In this paper, we develop a model of development and

intervention to estimate the implications of this trade-off for the age structure of optimal intervention.

To do this we present a generalised theory of state intervention in the development of children. We assume these interventions are directed at the reduction of educational under-performance and subsequent social exclusion. We take both the educational outcomes and the sort of intervention envisaged as given. The question is thus when or how often rather than what.

We cast the model in terms of human capital, which makes available the insights and techniques of the economic theory of capital. In this theory, capital is an enduring stock of goods, typically augmented by investment. This stock is productive and yields a return through time; the economic problem is to determine the level of investment that generates the maximum net value of these returns. Thus we define human capital to be any enduring human characteristic which generates a stream of benefits to the individual possessing the characteristic (or to others). Some human capital will be mental characteristics (mathematical ability for example), some not (good health, physical beauty). We are interested in varieties of adult human capital that are created or augmented by activities in childhood: we define these activities as investment in human capital. These definitions are very broad: most or all of the moral burden of parents in raising children can be interpreted in this format as facilitating investment in various forms of human capital.

We focus on mental human capital. In our context, this is to be understood as knowledge or habits of mind, interior to the individual, which influence or determine outcomes in adulthood. Typically human capital is thought of as those mental abilities that generate a return in the labour market. Thus the time and effort spent learning the times-tables in childhood creates a stock of facts which reaps the return through life of facility in arithmetical calculation, likely to be rewarded in the market-place. However such qualities as patience (Becker & Mulligan, 1997; Frederick, Lowenstein & O'Donoghue, 2002; Fuchs, 1992) and self-regulation (Alexander, Entwisle, & Dauber, 1993; Olson, Sameroff, Kerr, Lopez, & Wellman, 1999; Raver, Smith-Donald, Hayes, & Jones, 2005) may well be inculcated in childhood to yield a return outside the market-place at least as great as within it. For the purposes of the model developed below all we need to assume is that the behaviour flowing from a specific form and quantity of human capital can be measured and valued. An intervention is then interpreted as effecting an investment in this form of human capital, occurring when the natural, non-intervention rate of investment is calculated to be deficient. Just as in the conventional theory of capital, the economic problem is to determine the pattern of interventions that maximises the returns from changed behaviour net of the costs of the interventions.

In the terms of our investment approach, the State is in the position of valuing a capital asset - the child at adulthood - based on information currently available to it. If the forecast falls below a critical level, it will be efficient for the State to intervene to increase the rate of accumulation of human capital. If information is regularly available and used efficiently, then each forecast will be the best possible, given the information available. This implies that the series of forecasts evolves through time as

a random walk - a wandering series - as likely to rise as to fall[1]. This is precisely congruent to the behaviour of asset prices in the theory of capital markets. In our set-up, the forecast variable is the marginal benefit of human capital at adulthood. Intervention occurs when this exceeds the marginal cost of human capital at any given date, with the aim of increasing human capital until its marginal benefit equals its marginal cost. Thus the series of forecasts of marginal benefits wanders randomly in the region beneath the time-profile of marginal costs. This profile acts as a reflecting boundary, since the object of intervention is always to maintain the marginal benefit of human capital no greater than its marginal cost. This describes qualitatively the operation of an optimal system of intervention. We now fill in the details.

2 The model

We assume childhood lasts between $t = 0$ and $t = T$. At T some outcome is observed. In this paper we focus on test-scores and hence qualifications achieved as the key outcome of education, but other dimensions of behaviour such as criminality can be envisaged. We define human capital $x(t)$ at time (age) t as the expectation at t of the outcome at T, using all information available at t. This definition is a convenient way of making operational the notion of human capital. It would be more natural perhaps to define human capital by indices of intellectual and psychological development, but ultimately these would need to be mapped onto the behaviour at adulthood under consideration.

We assume human capital evolves according to the law

(1) $$dx(t) = dN(t) + \phi(t)i(t)dt$$

Here $dx(t)$ is the increment to human capital at time t. The variable $dN(t)$ is the instantaneous revision to the expectation of adult behaviour in the absence of intervention: this is the news that arrives at t about $x(t)$. We assume

(2) $$N(t) = \sigma(t)W(t)$$

where $W(t)$ is a Weiner process. A Weiner process is composed of standard normal increments which are independent at different dates[2]. This independence property is attractive because optimal forecasts at t of a given future event should embody all the information available at t and only change with the arrival of new information that is unforecastable at t. The variable $\sigma(t)$ scales the Weiner process to allow the variance of news to change over time.

The variable $i(t)$ is the State's expenditure on investment in the individual's human capital at time t. We shall assume there is a maximum feasible such investment:

$$0 \le i(t) \le i_m$$

---

[1] If, for example, the optimal forecast of a final outcome formed tomorrow were known to be larger than an optimal forecast formed today, then one could improve today's forecast by taking account of this. Formally, this is a consequence of the so-called Law of Iterated Projections.

[2] More exactly, a Weiner process is a stochastic variable in continuous time such that $W(t) - W(s)$ is Gaussian with variance $t - s$ and with independent non-overlapping increments.

The variable $\phi(t)$ measures the effectiveness of intervention at t. A single extra unit of expenditure $i \leq i_m$ at t will raise ultimate human capital $x(T) = x_T$ by $\phi(t)$: thus $\phi(t)$ is the marginal product of investment at time t. We define

$$\text{(3)} \qquad \kappa(t) = e^{\rho(T-t)} / \phi(t)$$

where ρ is the real interest rate (assumed constant). Then $\kappa(t)$ is the marginal cost evaluated at $t = T$ of a unit of x_T derived from expenditure on i at t. This schedule is a given of the problem and plays a central role in its solution. Note that this marginal cost is constant at each t while $i(t) \leq i_m$.

We assume the State derives £-utility $V(x_T)$ from an end-point x_T (valued at T). The State's problem is to maximise

$$\text{(4)} \qquad B = E_t \left[V(x_T) - \int_t^T i(s) e^{\rho(T-s)} ds \right]$$

at each date t by choice of a state-contingent path $i(s)$. "State-contingent" means here that $i(s)$ may be taken to respond to $x(s)$. The maximisation is performed subject to the law of motion given by (1).

3 The perfect foresight case

We consider first the case where the future is known with certainty i.e. there is never any news about the end state. In this case the planning problem is to maximise

$$\text{(5)} \qquad B = \left[V(x_T) - \int_0^T i(s) e^{\rho(T-s)} ds \right]$$

by choice of an investment plan $x(t)$. The solution to this problem has the following character, which may be proved by the Maximum Principle or by more elementary methods. To produce at minimum cost a quantity z of (intervention) human capital the State chooses an intervention set of the form

$$\Omega(\lambda) = \{ t \in [0,T]; \lambda \leq \kappa(t) \}$$

Then λ is chosen so that

$$\text{(6)} \qquad z = \int_{\Omega(\lambda)} i_m \phi(t) dt$$

Note that the investment level $i(t)$ is thus zero or the maximum i_m: a bang-bang solution. The value of λ that solves (6) is the marginal cost of increasing human capital at production level z. The relationship between z and λ given by (6) is thus a conventional marginal cost curve:

$$\text{(7)} \qquad MC = MC(z)$$
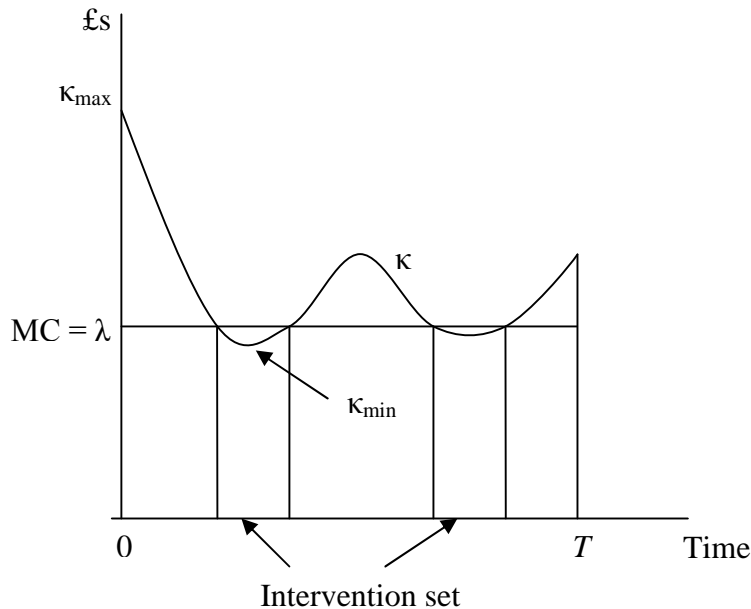
Figure 1 illustrates.

Figure 1 Intervention and the marginal cost curve

The curve is the instantaneous cost curve as defined above. The ordinate λ determines the intervention set as drawn. The intervention set then determines the production level of human capital. The value of λ is chosen so that this level of production delivers the desired z. When this is so, λ is marginal cost at this level of production. Note that when an individual enters or leaves a program of intervention, the instantaneous marginal cost of increasing x_T at this date given by $\kappa(t)$ is exactly the marginal cost of increasing x_T. At points of the intervention set other than on the boundary, overall marginal cost is greater than instantaneous marginal cost, but production cannot be expanded here because production is at a maximum. Note also that, while intervention is planned at $t = 0$, it does not begin immediately in Figure 1. As drawn, $\kappa(t)$ ranges between κ_{min} and κ_{max}; marginal cost κ_{min} corresponds to $z = 0$; marginal cost κ_{max} corresponds to maximum output where $i = i_m$ for all $[0, T]$.

The marginal cost curve partners the marginal benefit curve which has the form

(8) $$MB(z) = V_x(x(0) + z)$$

The optimal quantity of z is then determined by equating MC and MB in (7) and (8). Figure 2 illustrates.
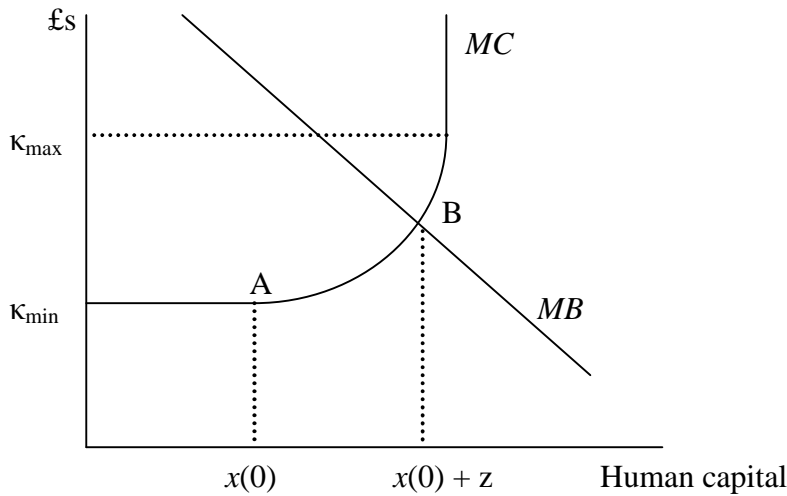
Figure 2 Determination of optimal level of intervention

Final level of human capital is at least $x(0)$. Marginal cost rises from the minimum level of κ_{min} at A. The optimal policy is determined by $MB = MC$. As drawn, this entails intervention to raise the final level of human capital to $x(0) + z$. If the MB were to pass to the left of point A in the diagram, then there is zero intervention. Thus:

(9)   Condition for intervention sometime $MB(x(0)) \geq \kappa_{min}$

Note that if MB passes above B then the intervention is at maximum level between 0 and *T* and the individual leaves the program with MB greater than MC.

If we assume that the profile does not vary across individuals then the cost-benefit analysis is completely conditioned by the initial level of human capital. Given this, the level of intervention and the corresponding MC are determined from Figure 2: this MC is then exported to Figure 1 where the intervention set is determined.

4 The effect of uncertainty

We now allow news about human capital to arrive over time. We define the value function $B(x,t)$ to be the value of having human capital level x at time t. This value is computed according to (4), assuming the (initial) level of human capital at t is x. For a small time increment one has

(10) $$B(x,t) = E_t \sup_{dz} \left[B(x + dz + \varepsilon, t + dt) - \kappa(t)dz \right]$$

where dz is investment in human capital and ε is the news arriving over the period $[t, t + dt]$. If the first term on the right is expanded as a Taylor series, one finds

(11) $$B(x,t) = E_t B(x + \varepsilon, t + dt) + dz[E_t B_x(x + \varepsilon, t + dt) - \kappa(t)]$$

where

$$dz = i_m \phi(t)dt \quad \text{when } E_t B_x(x+\varepsilon, t+dt) \geq \kappa(t)$$
$$= 0 \qquad \text{otherwise}$$

Taking the limit gives the condition for intervention at *t*:

(12) Condition for intervention $B_x(x,t) \geq \kappa(t)$

Taking the limit in (10) yields the Bellman equation, which in our context gives two partial differential equations for B, one when not intervening (in fact Cauchy's Heat equation) and a more complicated PDE when intervening. A boundary condition is given by the requirement that B not jump at the intervention boundary. These equations can be solved numerically for B and the intervention boundary. Our method of solution is to divide $[0,T]$ into a grid of subintervals and to base a backwards recursion on (11), which expresses the value function in terms of its expected future value and that of the derivative. The expected values can be calculated numerically, and the recursion set going at $t = T$ by the fact that

(13) $$B(x,T) = V(x)$$

5 Uncertainty and certainty equivalence

Uncertainty entails less intervention than the perfect-foresight case. One intuition for this result is that the possibility of *later* intervention establishes a bias in favour of non-intervention at any given time: at the margin one can correct unanticipated deterioration in human capital to some extent in the future, but unanticipated improvements involve a welfare loss since investment is irreversible. In the language of finance, the option of investing at any time has a positive value that disappears when the investment is made (see Dixit and Pindyck, 1994). For a slightly different intuition, consider the value function associated with no intervention at any time,

(14) $$B^0(x,t) = E_t V(x+\varepsilon)$$

where ε is the news that arrives after time *t*, so that $E_t \varepsilon = 0$. Define

(15) $$I = B - B^0$$

Then I is the incremental value of the system of intervention over non-intervention. One can think of I as the insurance value of the system: it offers protection if $x(t)$ deteriorates. One expects

(16) $$I_x < 0$$

since the insurance value will fall as x grows. From (14), given some regularity, one has $B_x^0 = E_t V_x(x+\varepsilon)$. Thus, since we take V to be quadratic, the expectation operator may be passed inside $V_x(x+\varepsilon)$ so that

(17) $$B_x^0 = V_x(x)$$

The condition for intervention (12) can now be written in terms of I as

(18) $$I_x \geq \kappa(t) - V_x(x)$$

When it is known that no news will arrive after time t, the intervention boundary is the locus of points (x,t) such that $\kappa(t) - V_x(x) = 0$. Call this the certainty-equivalence boundary. If $\kappa(t)$ is an increasing function of time, then the CE boundary is decreasing in time. Taking (16) and (18) together, one sees that intervention will not occur at the CE boundary. The marginal benefit from intervention is $I_x + V_x$, which is to be equated to κ: this is less than V_x, since an increment to x entails loss of value in the insurance component.

Figure 3 shows the two boundaries calculated as described above for the model we shall take as our benchmark below. Units are chosen for human capital so that the population mean is zero (in the absence of all intervention) and a single unit corresponds to innovation variance at $t = 0$. As time advances, human capital evolves in discrete time as a random walk from left to right in the diagram. Intervention occurs when the random walk crosses the blue curve. The vertical difference between the two curves represents extra caution induced by uncertainty.
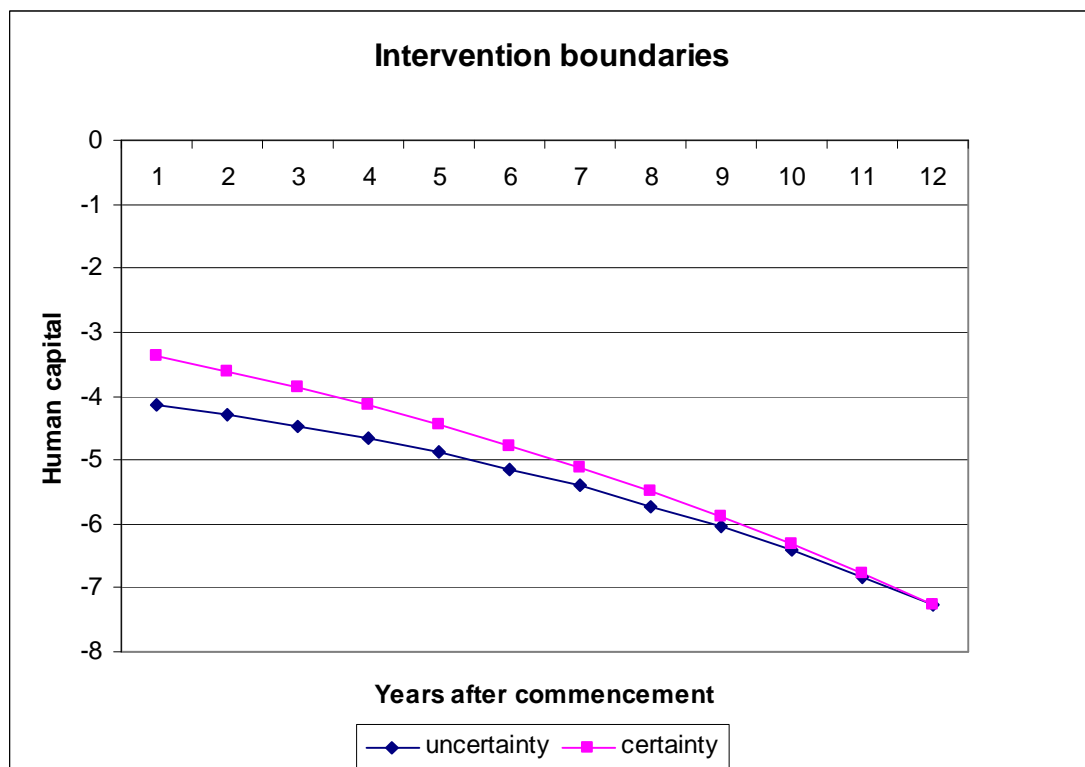

Figure 3 Intervention boundaries corresponding to uncertainty and certainty

6 Optimal mistakes

It is of considerable interest to know the extent of the mistakes made by optimal policy in this model. By mistakes here we mean firstly interventions that were not cost-effective in the sense that the value of the benefits turned out to be less than the cost, and secondly cases where interventions would have generated net benefits but were not made. These mistakes are mistakes only in hindsight and occur when the child's development takes an unpredictable turn. We emphasise that these mistakes are made under the action of an *optimal* policy, one that makes best use of information available at the time when the decision was made. The extent of such mistakes can be reduced only by improving the information available to the decision-maker. One might call them optimal mistakes.

One has a four-way classification of interventions, according to whether the intervention was made or not made, and whether it was cost-effective or not. See the table.

	Benefit $\geq$ Cost	Benefit $<$ Cost
Intervention	True positive	False positive
No intervention	False negative	True negative

Along the diagonal, true positives and true negatives refer to policy that was cost-effective, even with hindsight. In these cases, *ex post* the benefits were greater than the costs in cases where interventions were made *vice versa* when they were not. False negatives refer to cases where, at adulthood, some intervention in the past would have been cost-effective, but was not made, and *vice versa* for false positives.

7 Parameterization of the model

Our aim is to simulate the model which entails specifying the benefits and costs of intervention, as well as the stochastic structure of forecasts of outcomes..

7.1 Benefits We choose as unit of value the benefit at T from raising an individual from $x_T = \bar{x}_T - 2\sigma_x$ to $\bar{x}_T$ where $\bar{x}_T$ is the population average value of x_T and σ_x is the population standard deviation. We set $\bar{x}_T = 0$ and assume $MB(\bar{x}_T) = 0$ i.e. the planner values increments to human capital above the mean as zero. The benefit function is then

$$V(x_T) = -\frac{x_T^2}{4\sigma_x^2}$$

MB is obtained by differentiating *V* with respect to x_T.

7.2 Cost-structure We need to specify the intervention effectiveness function. We assume

$$\phi(t) = \phi(0)e^{-\gamma t}$$

where γ is a positive constant. Thus we parameterize $\phi(t)$ by its initial value $\phi(0)$ and its rate of decline γ. The cost-structure is then completed by ascribing a value to maximum intervention expenditure i_m.

We have found an alternative parameterization to be of use. Let θ be the cost of maximum intervention from 0 to T (thus implicitly the proportion of such cost to the value numeraire introduced above) and let τ be the number of standard deviations increase in x_T that such intervention would bring. Let ρ be the real interest rate. Then one has

$$\int_0^T i_m e^{\rho(T-t)} dt = \theta$$

and

$$\int_0^T i_m \phi(0) e^{-\gamma t} dt = \tau \sigma_x$$

which solve to give

$$i_m = \frac{\theta \rho}{e^{\rho t} - 1}$$

and

$$\phi(0) = \frac{\tau \sigma_x \gamma (e^{\rho T} - 1)}{\theta \rho (1 - e^{-\gamma T})}$$

For given values of ρ and γ, the cost-structure is thus determined by nomination of either the θ, τ pair or the $\phi(0), i_m$ pair. The latter are the more fundamental parameters, but the former are more interpretable and useful in determining a benchmark model. The average cost c_{SD} of a unit standard deviation increase in x_T from full intervention is

$$c_{SD} = \theta / \tau$$

Using the linearity of the MB schedule and the unit of value assumption from 7.1, one can show that, under certainty, one would intervene fully throughout childhood for all individuals with non-intervention human capital less than

$$x_T = -\sigma_x (2c_{SD} + 1)$$

Thus the system-parameter c_{SD} is the key cost in determining the prevalence of intervention. If $c_{SD} = 1/2$ it will be cost-effective to intervene fully for the lowest

2.5% of the population (with a normal distribution of x_T). Some interventions have low c_{SD} (presumably the proponents of free-milk in schools would argue this) and some have high c_{SD} (placing children in care).

7.3 Stochastic structure We have defined human capital $x(t)$ at time t as the expectation of adult behaviour formed at t. Let σ_0^2 be the population variance of $x(0)$. The innovation in the stochastic process $x(t)$ has variance $\sigma^2(t)$. We shall assume

$$\sigma^2(t) = \sigma^2(0)e^{-vt}$$

where v is a constant, and choose units of measurement so that $\sigma_0^2 = 1$. Note that the unconditional population variance is related to the variance of news by

$$\sigma_x^2 = \int_0^T \sigma^2(t)dt + \sigma_0^2 .$$

7.4 Summary of calibration requirements A simulation requires nomination of the six parameters in the table below.

Cost of full intervention relative to benefit numeraire	θ
Number of std devns increase from full intervention	τ
Rate of decline of intervention effectiveness	γ
Real interest rate	ρ
Initial variance of $x(t)$ in population	σ_0^2
Rate of decline of news-variance	v

The parameters θ and τ may be replaced by $\phi(0)$ and i_m for an alternative parameterization.

8 Benchmark assumptions

We formulate a benchmark set of parameters, around which deviations can be studied. The benchmark thus specifies the broad class of interventions we wish to consider in detail. This broad class is characterized by

- full childhood-long interventions are quite rare
- full intervention will not raise a typical recipient to the population average

In line with the discussion in 7.2, we take $\theta = 0.5$ and $\tau = 1$: this ensures that the most a full intervention can do is to raise x_T by one standard deviation, and that the implicit cost is such that about 2.5% of the population would be put on a childhood-long program from day one.

In this model whether intervention occurs early or late is determined by whether the slope of the marginal cost curve $\kappa(t)$ is positive or negative (respectively). This slope is positive if the rate of decrease γ in the effectiveness of intervention $\phi(t)$ is greater than the real interest rate ρ. It is commonly argued that the effectiveness of intervention declines rapidly as the child matures, so in the benchmark we have chosen $\gamma = 0.1$ and $\rho = 0.03$. Thus effectiveness approximately halves twice over the

course of childhood in the benchmark. The effect of a rising $\kappa(t)$ curve is that it is always cheaper to secure a given increase in final human capital by an intervention as early as possible. It follows that, under a regime of optimal intervention, late intervention occurs as a response only to an *unanticipated* deterioration in forecast final human capital.

To obtain orders of magnitude for the stochastic parameters, we estimated prediction equations for an index of academic performance at age 16 from the 1958 cohort of the National Child Development Survey. See Table 1.

Table 1: Change in error variance of outcome prediction as children mature

	Predictors of performance at age 16		
Prediction at age	Family type	Test scores	Prediction MSE
0 years	Yes	No	1.87
7 years	Yes	Yes	1.31
11 years	Yes	Yes	1.04

Between 0 and 7, and between 7 and 11, the prediction error declines by about 0.08 units per annum and the implied value of σ_0^2 (the variance of $x(0)$ in the population) is about 12.5 in units of the innovation variance. As a benchmark we assume that the news variances decline at 1% per annum and that the variance $\sigma_0^2 = Var(E_0 x_T)$ is 12.5.

9 Simulation of the benchmark

We shall study the properties of the model by simulation. A hypothetical individual at time $t = 0$ is given an initial human capital which is a random variable of variance σ_0^2. Subsequently the individual's human capital receives random shocks of variance $\sigma^2(t)$ and intervention takes place according to the optimal criteria outlined above. This path through childhood is then assessed in terms of costs and benefits. This operation is repeated a large number of times (100,000) which enables calculation of system characteristics such as benefit-cost ratios and participation rates.

We have simulated this model for the design:

τ	γ	ρ	σ_0^2	$\nu\theta$
0.5	1.0	0.1	0.03	12.5	0.01

We filter out all interventions of expected duration less than one year. We obtain a benefit-cost ratio of 1.58. At one time or another, 15.6% of the population are on a program; the total reduction in inequality (measured as the standard deviation of final human capital) is 7.4%. Total population participation is shown below.

Participation in programs
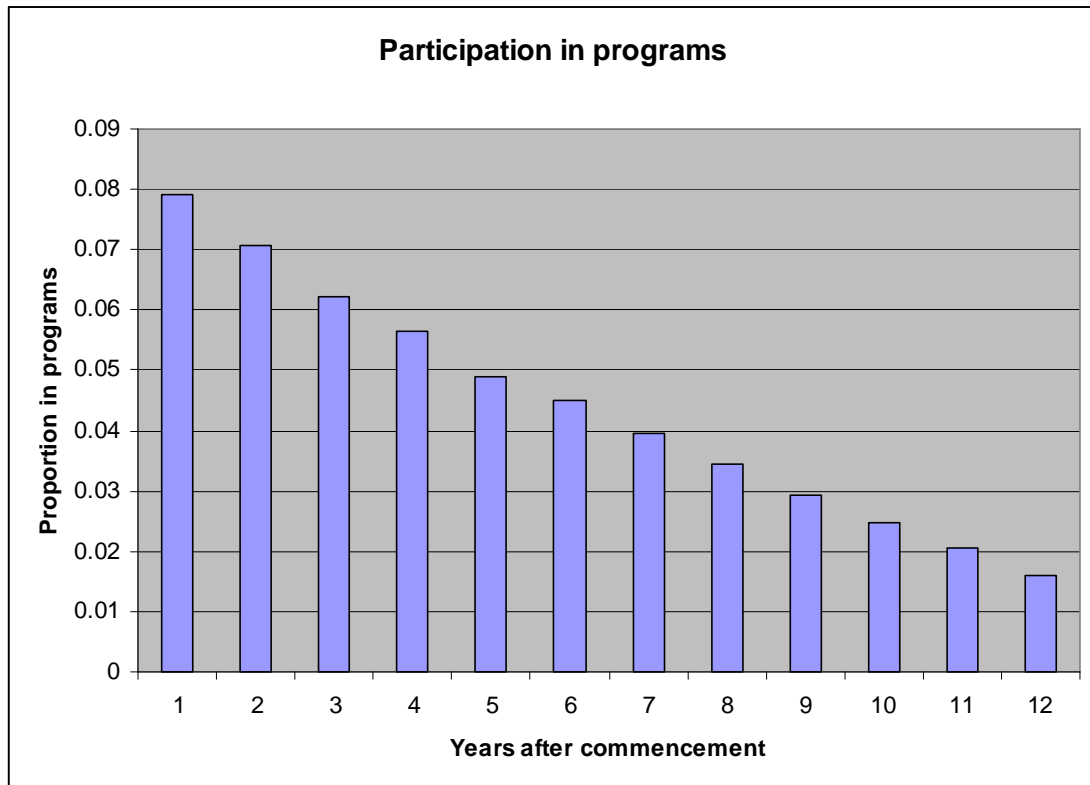
Just under 8% of the population are placed on a program immediately. The proportion then declines to about 1.5% after 12 years.

The frequency distribution of the total length of time children spend on programs is given below.

Frequency of treatment durations
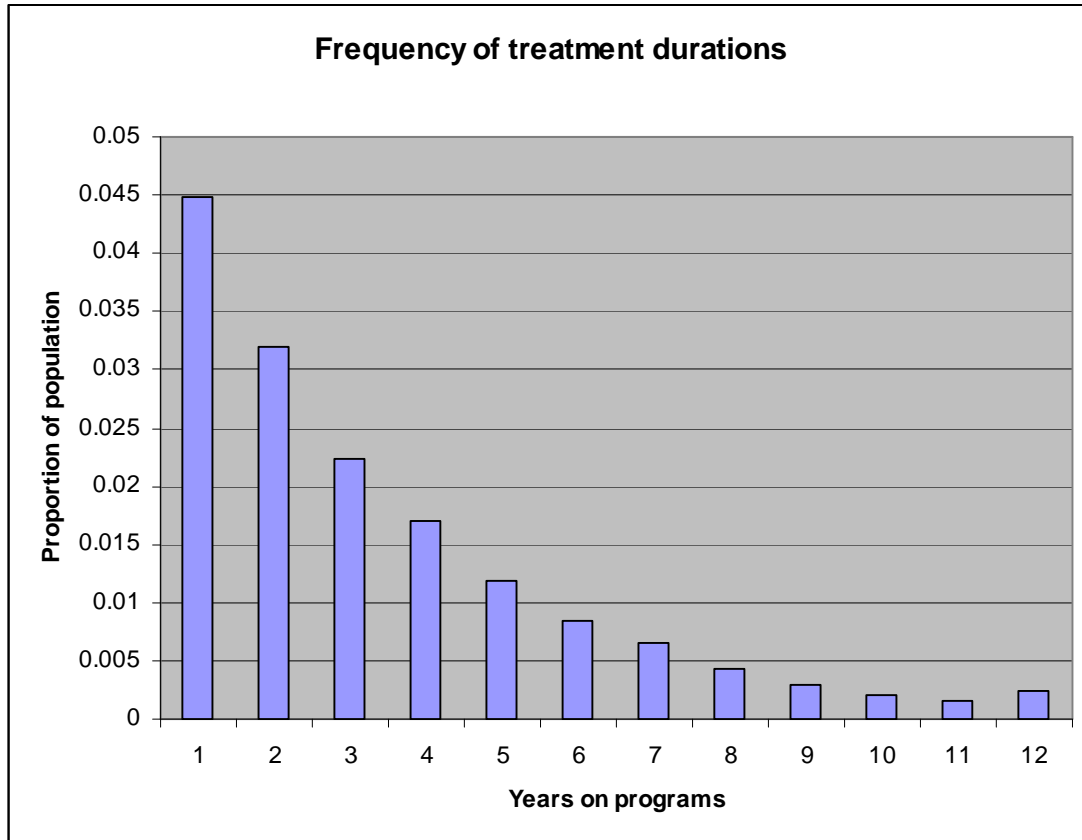
The most common time spent on a program is one year. The frequency falls steadily with duration. About 1% of the population spend nine or more years on programs.

Accuracy of diagnosis is summarised in the table below.

	Benefit ≥ Cost	Benefit < Cost
Intervention	0.112	0.044
No intervention	0.111	0.733

In the true-positives cell, 11.2% of the population receive a treatment that is cost-effective *ex post*, while, in the true-negatives cell, 73.3% receive no treatment when treatment would not be cost-effective[3]. Both false-positives (4.4%) and false-negatives (11.1%) are large, indicating that misdiagnosis is very common. This may seem paradoxical at first sight in an optimal system of intervention, but it is an inevitable feature of a model of this sort. At time $t = 0$ all individuals with poor forecasts will be placed on a program. Of the others, they are not considered at-risk for the present; subsequently the forecasts of their final level of human capital evolve as a random walk as described above. Intervention takes place only when this random walk crosses the reflecting boundary in Figure 3. Thus all new interventions after day

---

[3] In our framework treatment always benefits the individual: cost-effective treatment requires the benefits *as measured by the planner*.

one are more or less borderline cases, and misdiagnosis *ex-post* can be expected to be common.

10 Sensitivity

Table 2 presents results of simulation for various values of the rate of decline of intervention-effectiveness.

Table 2: Simulations varying the rate of decline of intervention-effectiveness

	Benchmark but with $\gamma = 0.05$	Benchmark ($\gamma = 0.10$)	Benchmark but with $\gamma = 0.15$
Benefit-cost ratio	1.663	1.578	1.647
Prop ever on program	0.230	0.156	0.121
Prop on day one	0.045	0.078	0.088
Prop at 18 years	0.100	0.016	0.000
Long-term treatments	0.022	0.009	0.001
False positives	0.059	0.046	0.034
Reduction in inequality	0.131	0.074	0.045

Note: In the benchmark model, σ_0^2 = initial variance of $x(t)$ in the population = 12.5; γ = rate of decline of intervention effectiveness = 0.10; ν = rate of decline of news variance = 0.01; κ = units of std. dev. above average at which MB is 0 = 0 θ = cost of full intervention relative to benefit *numeraire* **= 0.5.**

In the first column we assume that effectiveness declines by 5% per annum, thus approximately halving over the course of childhood. This results in a scheme of larger scope than the benchmark (23% of the population are on a program at one time or another compared to 15.6% under the benchmark). Note that intervention tends to occur late rather than early, in contrast to the benchmark. In a world of certainty, intervention occurs early or late according as the rate of decline of effectiveness is greater than or less than the real interest rate. With $\gamma = 0.05$, the rate of decline is greater than $\rho = 0.03$, but the presence of uncertainty provides another reason to intervene late which turns out to be stronger than the effect of the real interest rate. This example shows the early/late question depends crucially on the cost technology of intervention.

In the third column we increase the rate of decline of intervention-effectiveness from 0.15, so that effectiveness at the end of childhood is approximately one eighth of its initial value. This essentially wipes out later intervention: there are no treatments at age 18 and only 0.1% of the population receive long-term treatments.

Note that the proportion of the population on a program at day one increases with the rate of decline of intervention effectiveness (row three). On day one the effectiveness of intervention is the same in all three columns (because $\phi(0)$ is the same). When later intervention is costly, however, the planner has an added precautionary incentive to intervene early, in case human capital unexpectedly deteriorates.

In Table 3 we present simulations for different values of the initial variance of human capital, holding everything else as in the benchmark. Naturally one expects the higher the variance of initial human capital, the higher the level of intervention at all ages,

and the greater the reduction in inequality. This is borne out: all measures of intervention increase monotonically with σ_0^2, as does the reduction in inequality. Intervention is early rather than late in the last three columns, but for $\sigma_0^2 = 4$ one finds that late intervention is more prevalent. The logic behind this is that if the variance of initial human capital is small, so that one has little information about children's prospects, then the best forecast of adult behaviour is average adult behaviour, which does not warrant intervention. With the passage of time, some children will experience falls in observed human capital and be subject to intervention. Thus intervention will increase over time. If, however, the rate of decline in intervention-effectiveness is high enough, late intervention will be low. In these circumstances the path of average intervention levels will not be monotonic: intervention will be low both for young children and old children, with a maximum intervention level somewhere in between. In the model in the first column of Table 3, intervention is at its maximum seven years after commencement i.e. halfway through childhood.

Table 3: Simulations varying the initial variance of human capital, σ_0^2

	Benchmark but with initial variance=4	Benchmark but with initial variance=10	Benchmark (initial variance=12.5)	Benchmark but with initial variance=15
Benefit-cost ratio	1.383	1.521	1.578	1.638
Prop ever on program	0.064	0.134	0.156	0.179
Prop on day one	0.007	0.057	0.078	0.100
Prop at 18 years	0.009	0.013	0.016	0.019
Long-term treatments	0.001	0.006	0.009	0.013
False positives	0.020	0.041	0.046	0.050
Reduction in inequality	0.027	0.062	0.074	0.085

Table 4 presents results for different values of the cost of full intervention relative to the benefit *numeraire*: one sees that the general prevalence of intervention depends very strongly on the cost of intervention. With intervention 20% less costly than in the benchmark, over one in five of the population participates in programs, whereas, with intervention 20% more costly than the benchmark, the proportion falls to about one in ten.

Table 4: Simulations for different values of the cost of full intervention θ

	Benchmark but with cost of intervention 20% smaller	Benchmark	Benchmark but with cost of intervention 20% larger
Benefit-cost ratio	1.731	1.578	1.490
Prop ever on program	0.218	0.156	0.110
Prop on day one	0.105	0.078	0.058
Prop at 18 years	0.035	0.016	0.006
Long-term treatments	0.019	0.009	0.004
False positives	0.063	0.046	0.032
Reduction in inequality	0.104	0.074	0.051

In Table 5 we give simulations for different effectiveness levels (quantity measure) of intervention. Proportional increases in the effectiveness of intervention have similar but not equivalent effects to corresponding reductions in costs in Table 4. In both cases there is no indication that early intervention would be overturned by plausible variation in these parameters.

Table 5: Simulations for different values of the effectiveness of full intervention τ

	Benchmark but with effect of full intervention 20% smaller	Benchmark	Benchmark but with effect of full intervention 20% larger
Benefit-cost ratio	1.462	1.578	1.707
Prop ever on program	0.106	0.156	0.202
Prop on day one	0.058	0.078	0.092
Prop at 18 years	0.007	0.016	0.027
Long-term treatments	0.005	0.009	0.012
False positives	0.032	0.046	0.056
Reduction in inequality	0.043	0.074	0.105

Table 6 reports simulations varying the real interest rate and the rate of decline in innovation-variance.

Table 6: Simulations varying the real interest rate and the rate of decline in innovation-variance

	Benchmark but with $\rho = 0.06$	Benchmark but with $v = 0.1$
Benefit-cost ratio	1.428	1.609
Prop ever on program	0.095	0.145
Prop on day one	0.029	0.084
Prop at 18 years	0.023	0.005
Long-term treatments	0.006	0.007
False positives	0.026	0.035
Reduction in inequality	0.047	0.082

High real interest rates are favourable to later expenditures and this is reflected in the first column where we see late interventions are almost as large as early interventions. In the second column we allow innovations to human capital to become smaller with the passage of time, thus enabling more accurate forecasts of adult outcomes. Compared with the benchmark, the benefit-cost ratio is higher, for a somewhat smaller program size. The accuracy-of-diagnosis matrix is:

	Benefit ≥ Cost	Benefit < Cost
Intervention	0.111	0.035
No intervention	0.096	0.757

This matrix should be compared to the analogous matrix in Section 9. One sees that the off-diagonal entries are smaller, reflecting fewer errors in assignments to

programs, both of those who shouldn't have been on a program but were, and who should have been but weren't. This is a return to more accurate forecasts.

11 Extensions

One of the main concerns of practitioners is the negative effect on individuals from being placed in an intervention category, sometimes called *stigma-effect* (see e.g. Foster et al., 1972; Horowitz & Garber, 2001; Perlick et al., 1991) When such effects are important, there will be fewer interventions. Problems of measurement aside, it is straightforward to incorporate this in our framework. Figure 4 is a variant of Figure 2 in which the initial level of human capital $x(0)$ at point A is chosen so that the area between the MC and MB curves is exactly the stigma cost. Since this area is the total non-stigma benefit from an intervention, interventions are efficient only for initial levels of human capital to the left of A. In contrast, without stigma effects, some intervention would occur for all initial levels of human capital to the left of point B.
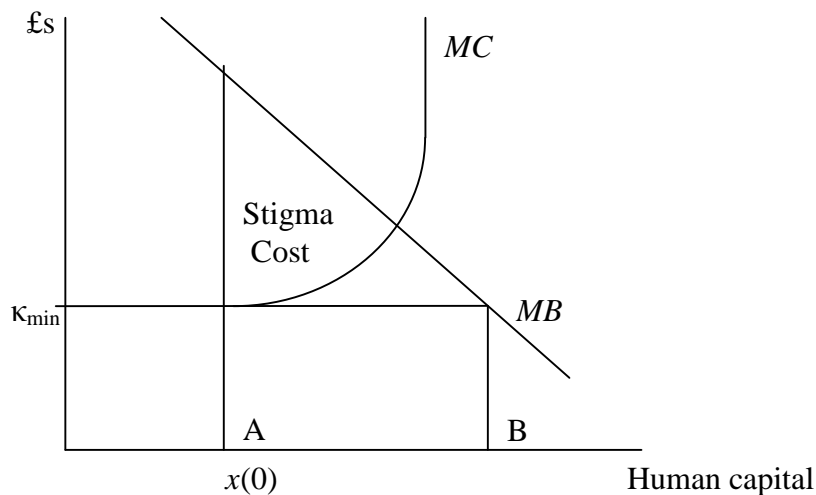

Figure 4 Reduced interventions with stigma costs

Another extension of the framework would address the relative trade-offs between offering effective help at the lowest end of the human capital distribution and at low but less extreme points. Some argue that current policy effectively abandons the lowest achieving tranche of children and concentrates on helping the next tranche (Gillborn & Youdell, 2000). In the model as developed above, the law of motion assumes that a given quantity of investment raises human capital by a given amount, irrespective of the initial level. A plausible alternative in line with Heckman's notion of complementarities between learning gained at one point in time and later learning would be that a given level of investment would raise human capital by a given proportional amount. In this case, the law of motion would be

$$dx(t) / x = dN(t) + \phi(t)i(t)dt$$

and the marginal cost of a unit of human capital derived from intervention would be

$$MC = \kappa(t)/x$$

where $\kappa(t)$ is the original marginal cost function.[4] The MC = MB region of (x,t) space is shown in Figure 5. There are two intervention boundaries, an upper and a lower, and intervention occurs only within the convex region enclosed by them. If the decline in intervention-effectiveness $\phi(t)$ is modest then at the end of childhood at point A there will still be some individuals for whom intervention is appropriate. If however the decline in intervention-effectiveness is sufficiently great, then all intervention will cease before childhood ends, in this case represented by point B. Despite the fact that the benefits of raising the human capital of those below the lower intervention boundary might be very large, the costs of doing so are even greater than the benefits.
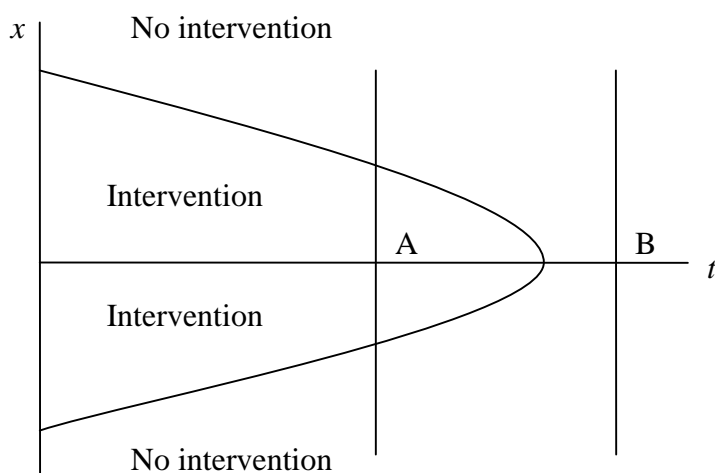

Figure 5 Intervention boundaries when the marginal product of intervention depends on human capital

It follows that the qualitative character of a system of intervention depends crucially on the relationship between the effectiveness of intervention and existing levels of human capital. Very little is known about this relationship. It is also possible that the rate of return on investments in human capital will be highest for those with the lowest current levels as many of these individuals may have received low levels of private (family) investment and so may benefit from high returns to investment in line with the law of diminishing returns.

12 Summary and conclusions
We have developed a general model of optimal state-intervention in childhood that takes account of the presence of uncertainty in forecasts of adult outcomes, as well changes in the effectiveness of intervention at different ages. The model requires parameters to describe:

---

[4] Note that under this formulation one would require the index of human capital to be always positive.

- The preferences of the state planner
- The cost-structure and effectiveness of intervention
- The stochastic structure of forecasts of adult outcomes

Preferences of the planner. A £-value needs to be placed on increased human capital. We have chosen as numeraire the value of raising human capital from two standard deviations below the mean to the mean.

Cost-structure and effectiveness of intervention. The model requires nomination of the cost of full (childhood-long) intervention relative to the benefit numeraire. This quantity is thus the £-cost of full intervention divided by the £-value of raising an individual by two standard deviations to the mean. In our benchmark model we have chosen this parameter to be 0.5, which implies that intervention to so raise the human capital of an individual at the bottom of the distribution would be well worth the cost, if it were possible. A second required parameter is the number of standard deviations of increase in human capital that full intervention can achieve. We have chosen this to be unity. These two assumptions characterize the benchmark as a system where full intervention throughout childhood occurs but is rare, and where raising one of these full-intervention children to the mean typically does not occur. A third required parameter is the rate of decline of intervention effectiveness. The qualitative nature of the program of intervention depends heavily on this parameter. We have chosen it in the benchmark model to be 0.1 (compared to a real interest rate of 0.03). This implies that intervention at the end of childhood is about a quarter as effective as it was at the beginning.

Stochastic structure of forecasts of adult outcomes. The model requires both the population variance of the day-one forecasts of adult outcomes and the year-on-year decline of the error variance of forecasts. Examination of forecast equations from the 1958 NCDS cohort suggested that the variance of the year-on-year forecasts declined by about the same amount each year, and that the initial variance was about 12.5 that number. We chose these magnitudes for our benchmark model, but we allowed the year-on-year forecast errors to decline by a token 1% per annum.

For this specification of parameter values, we find that quite an activist policy of intervention is optimal. About 8% of children are placed on a program immediately at five years. Participation in programs declines steadily to about 1.5% by age 17. The most common duration of program is one year, for about 4.5% of the population. About 1% of the population spend nine or more years on programs. Misdiagnosis *ex-post* is quite common, despite the optimality of the procedure. About 11.1% of the population experience no intervention when this would have been cost-effective *ex-post*, while 4.4% experience intervention that is not cost-effective *ex-post*.

Sensitivity We have subjected all our parametric assumptions to sensitivity analysis. Changes in the parameters will in general change the predictions of the model. Intervention effectiveness and the cost of intervention act similarly but in opposite directions to determine the general scope of intervention as seen in Tables 4 and 5. If the rate of decline of intervention-effectiveness is increased substantially from our benchmark of 10% then late-intervention is effectively eliminated. For moderate rates of decline, however, late intervention can be more prevalent than early intervention. Similarly, high real interest rates can favour late intervention. We have found as well

that, if the initial variance of human capital in the population is low, so that early forecasts of adult outcomes tend to be poor, then it can happen that the prevalence of optimal intervention will rise over time, only to fall towards the end of childhood.

Early versus late A major aim was to cast some light on the early versus late intervention debate. We have formulated a benchmark model to have the properties that full intervention is fairly rare, only of limited effectiveness, and declines in effectiveness over the course of childhood to about one quarter of its value at the beginning. For a stochastic structure parallel to the information in NCDS data, we have found that intervention is unambiguously early rather than late. However there are models not too far distant in parameter space where this is reversed. Perhaps the most critical is the rate of decline of intervention-effectiveness. If this is 5% per annum rather than 10% per annum, then intervention is late rather than early.

Informational requirements Policy-makers contemplating a program of intervention need to consider the three sets of parameters indicated above. In particular, they need to quantify:

- The social benefits of moving an individual out of the lower tail of the human capital distribution.
- The costs and effectiveness of intervention. Costs are easy, but effectiveness is harder. One needs to know the age-profile of intervention-effectiveness. One needs to know the effectiveness of intervention for those with very low human capital.
- The accuracy of forecasts of adult outcomes. If the only available forecasts are weak predictors of adult outcomes then the scope for policy is reduced. We have found that forecasts based on the NCDS are sufficiently accurate to support a fairly activist policy of intervention in education, but this may not be so in other contexts.

References

Alexander, K.L., Entwisle, D.R. and Dauber, S.L. (1993) 'First grade classroom behavior: its short- and long-term consequences for school performance', *Child Development*, 64: 801-814.

Bandura, A. (1994). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.

Becker, G.S. and C.B. Mulligan. "The Endogenous Determination of Time Preference." Quarterly *Journal of Economics* 112 (1997): 729-58.

Dixit, A. K., and R. S. Pindyck. *Investment Under Uncertainty*. Princeton NJ: Princeton University Press, 1994.

Feinstein (2006). "Predicting adult life outcomes from earlier signals: modelling pathways through childhood." WBL Discussion Paper, Institute of Education, London.

Foster, J., Dinitz, S., Reckless, W. (1972) "Perceptions of Stigma following Public Intervention for Delinquent Behavior." *Social Problems*, Vol. 20, No. 2 (Autumn, 1972), pp. 202-209

Frederick, S; Lowenstein, G; O'Donoghue, T. "Time Discounting and Time Preference: A Critical Review." *Journal of Economic Literature* 40 (2002): 351-401.

Fuchs, Victor R., (2004). "Reflections on the Socio-Economic Correlates of Health." *Journal of Health Economics* 23: 653-61.

Fuchs, Victor R., (1982) "Time Preference and Health: An Exploratory Study." *Economic Aspects of Health*. Ed. V.R. Fuchs. Chicago: University of Chicago Press. 93-120.

Gillborn, D., & Youdell, D., (2000) *Rationing Education*. Buckingham, Open University Press.

Goswami, U. (2007). *Cognitive Development: The Learning Brain*. Taylor & Francis (Psychology Press).

Heckman J. (2003) "Human capital policy." In Heckman J, Krueger A (Eds.), *Inequality in America: What role for human capital policy?* Cambridge, MA: MIT Press.

Horowitz, J., & Garber, J., (2006). "The Prevention of Depressive Symptoms in Children and Adolescents: A Meta-Analytic Review." *Journal of Consulting and Clinical Psychology*. Vol. 74, No. 3, 401–415

Olson, S.L., Sameroff, A., Kerr, D.C.R., Lopez, N.L. and Wellman, H.M. (2005) 'Developmental foundations of externalizing problems in young children: the role of effortful control', *Development and Psychopathology*, 17: 25-45.

Perlick, D., Rosenheck, R., Clarkin, J., Sirey, J., Jamelah Salahi, B.S., Struening, E. and Link, B. (2001) "Stigma as a Barrier to Recovery: Adverse Effects of Perceived Stigma on Social Adaptation of Persons Diagnosed With Bipolar Affective Disorder." *Psychiatric Services*. 52:1627-1632, Dec.

Piatt, W. (2003) "Social Mobility." *New Economy* 10 (4), 187–188.

Raver, C.C., Smith-Donald, R., Hayes, T. and Jones, S.M. (2005) "Self-regulation across differing risk and socio-cultural contexts: preliminary findings from the Chicago School Readiness Project." Paper presented at the biennial meeting of the Society for Research in Child Development, Atlanta, GA.

Reynolds, A. J., Wang, M. C., & Walberg, H. J. (Eds.). (2003). *Early childhood programs for a new century.* Washington, DC: Child Welfare League of America

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25, 68-81.